


```

RESULT 2
E311_ADE06 STANDARD: PRT: 101 AA.
ID E311_ADE06
AC 055653:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Early E3A 11.6 kDa glycoprotein.
OS Human adenovirus type 6.
VIRUSES: dsDNA viruses, no RNA stage: Adenoviridae; Mastadenovirus.
NCBI_TaxID=10534;
RX [1]
RN SEQUENCE FROM N.A.
RA Reischmann H., Scharschmidt E., Geisler B., Hausmann J., Ortmann D.,
RA Bauer U., Flunker G., Seidel W.;
RL SUBMITTED (DEC-1997) TO THE EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
CC KDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y16037; CA75991.1;
CC DR Early protein: Glycoprotein; Transmembrane.
CC KW TRANSMEM 41 62
CC FT CAROHD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 101 AA; 11613 MW; EP219000939E3B4B CRC64;

Query Match 40.2%; Score 35; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 14e-30;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 MMFSLALMFVCLIMLIMLICLKRRAPIYRPI 75
Db 41 MMFSLALMFVCLIMLIMLICLKRRAPIYRPI 75

RESULT 3
E311_ADE05 STANDARD: PRT: 93 AA.
ID E311_ADE05
AC P17590:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Early E3A 10.5 kDa glycoprotein.
OS Human adenovirus type 5.
VIRUSES: dsDNA viruses, no RNA stage: Adenoviridae; Mastadenovirus.
NCBI_TaxID=28285;
RX [1]
RN SEQUENCE FROM N.A.
RA MEDLINE=85092388; PubMed=2981456;
RA Claderas C., Wold W.S.M.;
RT "DNA sequence of the early E3 transcription unit of adenovirus 5.";
RL Virology 140:28-43(1985).
RN [2]
RN COMPLETE GENOME.
RX MEDLINE=92087470; PubMed=1727603;
RX Chroboczek J., Bieber F., Jacot B.;
RT "The sequence of the genome of adenovirus type 5 and its comparison
RT with the genome of adenovirus type 2.";
RL Virology 186:280-285(1992).
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED.

```

```

CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
CC KDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M73260; NOT_ANNOTATED_CDS.
CC DR EMBL: X03002; CA26784.1;
CC DR PIR: A05245; ERAD53.
CC KW Early protein: Glycoprotein; Transmembrane.
CC FT TRANSMEM 34 55
CC FT CAROHD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 93 AA; 10523 MW; 008AD087AAB45ABF CRC64;

Query Match 27.6%; Score 24; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 8.7e-19;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 MMFSLALMFVCLIMLIMLICLKR 64
Db 34 MMFSLALMFVCLIMLIMLICLKR 57

RESULT 4
HRK_MOUSE STANDARD: PRT: 92 AA.
ID HRK_MOUSE
AC P70678:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Activator of apoptosis harakiri (neuronal death protein dp5).
GN HRK OR DP5 OR BID3.
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090, 10116;
RN [1]
RN SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Imaizumi K.;
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ACTIVATES APOPTOSIS AND INTERACTS SELECTIVELY WITH
CC SURVIVAL-PROMOTING PROTEINS BCL-2 AND BCL-XL (BY SIMILARITY).
CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR THE INDUCTION OF CELL
CC DEATH (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D83697; BAA12065.1;
CC DR EMBL: D83698; BAA12066.1;
CC DR MGD: MGI:1201608; B1d3.
CC DR InterPro: IPR000712; Bcl_2.
CC DR PROSITE: PS01259; BH3; 1.
CC FT APOPTOSIS.
CC SO SEQUENCE 92 AA; 10078 MW; 95760B9E8A88F1DE CRC64;

Query Match 6.9%; Score 6; DB 1; Length 92;

```

Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 RRRARP 69
|11111
DB 50 RRRARP 55

RESULT 5

VNBP_POPMV STANDARD; PRT; 121 AA.
ID VNBP_POPMV
AC 002123;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 01-OCT-1993 (Rel. 27, Last annotation update)
DE 14 kDa protein (Putative nucleic acid-binding protein).
OS Poplar mosaic virus (isolate ATCC PV275) (PMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.
OX NCBI_TaxID=31709;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-92333281; PubMed-1629709;
RA Henderson J., Gibbs M.J., Edwards M.L., Clarke V.A., Gardner K.A.,
RA Cooper J.I.;
RT "Partial nucleotide sequence of poplar mosaic virus RNA confirms its
RT classification as a Carlavirus.";
RL J. Gen. Virol. 73:1887-1890(1992).
CC -1- FUNCTION: MAY ACT AS A REGULATORY FACTOR DURING VIRAL
CC TRANSCRIPTION.
CC -1- SIMILARITY: TO NUCLEIC ACID-BINDING PROTEINS FROM OTHER
CC CARLAVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X65102; CAA46227.1; -
CC DR EMBL: D13364; BAA02629.1; -
CC DR PIR: J01646; J01646.
CC DR PIR: S23876; S23876.
CC DR InterPro: IPR002568; Carla_C4.
CC DR Pfam: PF01623; Carla_C4; 1.
CC KM Zinc-finger; DNA-binding.
CC FT ZN_FING 62 83
CC FT ZN_FING 62 83
CC SEQUENCE 121 AA; 14451 MW; 705205527BF7D3CF CRC64;

Query Match 6.9%; Score 6; DB 1; Length 121;

Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 KRRRAR 68
|11111
DB 52 KRRRAR 57

RESULT 6

NIFM_FRASE STANDARD; PRT; 126 AA.
ID NIFM_FRASE
AC 092570;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nitrogenase stabilizing/protective protein nifm.
GN NIFM.
OS Frankia sp. (strain EuIK1).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Frankineae; Frankiaceae; Frankia.
OX NCBI_TaxID=47227;

RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-EuIK1;
RA Chung-Sun A., Ji-Tae K., Won-Jin K., Won-Young Y.;
RT "Nif-gene organization and nucleotide sequences from Frankia EuIK1
RT strain.";
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: May protect the nitrogenase Fe-Mo protein from oxidative
CC damage (by similarity).
CC -1- SUBUNIT: Homotrimer; associates with nifD (by similarity).
CC -1- SIMILARITY: BELONGS TO THE NIFW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF19361; AAD17267.1; -
CC DR EMBL: AF19361; AAD17267.1; -
CC KM Nitrogen fixation.
SO SEQUENCE 126 AA; 14287 MW; 295D8C9102FD317E CRC64;

Query Match 6.9%; Score 6; DB 1; Length 126;

Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 TTATGL 20
|11111
DB 80 TTATGL 85

RESULT 7

YRN6_CAEEL STANDARD; PRT; 136 AA.
ID YRN6_CAEEL
AC 009420;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 14.2 kDa protein R0781.6 in chromosome X.
GN R0781.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Kerhaw J.;
RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z48621; CAA8543.1; -
CC DR WormPep: R0781.6; CE01632.
CC DR InterPro: IPR002601; C6.
CC DR Pfam: PF01681; C6; 1.
CC KM Hypothetical protein.
SO SEQUENCE 136 AA; 14169 MW; D7F47A83F3F81787 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 136;

Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 T1APT 10
111111
DB 31 T1APT 36

RESULT 8

YB9E_YEAST STANDARD; PRT; 142 AA.
ID YB9E_YEAST
AC P38337
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 16.4 kDa protein in POP4-SHM1 intergenic region.
GN YBR258C OR YBR1726.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE-9320397; PubMed-8465606;
RA Doignon F., Bileau N., Crouzet M., Algle M.;
RT "The complete sequence of a 19,482 bp segment located on the right
arm of chromosome II from Saccharomyces cerevisiae.";
RL Yeast 9:189-199(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X70529; CAA49922.1; -;
DR EMBL: Z36127; CA885221.1; -;
DR PIR: S32960; S32960.
DR SGD: S0000462; YBR258C.
KW Hypothetical protein.
SQ SEQUENCE 142 AA: 16429 MW: 12873FOC3DD25D7 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GLTSAL 24
111111
DB 71 GLTSAL 76

RESULT 9

RK34_SPTOL STANDARD; PRT; 152 AA.
ID RK34_SPTOL
AC R82244;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L34, chloroplast precursor.
GN RPL34.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 92-101, AND MASS SPECTROMETRY.
RC STRAIN-CV. ALVARO; TISSUE=leaf;
RX MEDLINE-20435798; PubMed-10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
the 50 S subunit of an organelle ribosome (chloroplast).";

RL J. Biol. Chem. 275:28466-28482(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- MASS SPECTROMETRY: MW-67.1; METHOD-Electrospray; RANGE-92-152.
CC -1- SIMILARITY: BELONGS TO THE L34 FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF238221; AAF6457.1; -;
DR InterPro: IPR000271; Ribosomal_L34.
DR Pfam: PF00468; Ribosomal_L34; 1.
DR PROSITE: PS00784; RIBOSOMAL_L34; FALSE NEG.
KW Ribosomal protein; Chloroplast; Transit peptide; rRNA-binding.
FT TRANSIT 1 91
FT CHAIN 92 152 50S RIBOSOMAL PROTEIN L34.
SQ SEQUENCE 152 AA: 16095 MW: 755A990D441ADB18 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 152;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 LKRRRA 67
111111
DB 127 LKRRRA 132

RESULT 10

HES2_MOUSE STANDARD; PRT; 157 AA.
ID HES2_MOUSE
AC O54792;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Transcription factor HES-2 (Hair cell enhancer of split 2).
GN HES2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129/J;
RX MEDLINE-98234545; PubMed-9570950;
RA Nishimura M., Isaka F., Ishibashi M., Tomita K., Tsuda H.,
RA Nakaniishi S., Kageyama R.;

RT "Structure, chromosomal locus, and promoter of mouse Hes2 gene, a
homologue of Drosophila hairy and enhancer of split.";
RL Genomics 49:69-75(1998).
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH
CC PROTEIN FOR THEIR TRANSCRIPTION.
CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
CC -1- SUBCELLULAR LOCATION: Nucleus.

CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
CC HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNG),
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A
CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
CC HAIR-RELATED PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AB009967; BAA24091.1; -
MGI:1098624; Hes2.
InterPro: IPR003015; HLH_MYC.
InterPro: IPR001092; HLH_dlm.
InterPro: IPR003650; Orange.
Pfam: PF00010; HLH; 1.
SMART: SM00353; HLH; 1.
SMART: SM00511; ORANGE; 1.
PROSITE: PS00038; HELIX_LOOP_HELIX; 1.
Transcription regulation; DNA-binding; Nuclear protein; Repressor.
DNA_BIND 13 26
FT DOMAIN 139 148
FT DOMAIN 154 157
WRPW MOTIF (REQUIRED FOR ACTIVITY)
(BY SIMILARITY).
SEQUENCE 157 AA: 17231 MW: 570A0C67F4992EA7 CRC64:

Query Match 6.9%; Score 6; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 KRRRAR 68
|||||
DB 23 KRRRAR 28

RESULT 11

HES2_RAT STANDARD; PRT: 157 AA.
AC P33429;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Transcription factor HES-2 (Hairy and enhancer of split 2).
HES2 OR HES-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Embryonic brain;
RX MEDLINE-9335886; PubMed-8354270;
RA Ishibashi M., Sasai Y., Nakanishi S., Kageyama R.;
RT Molecular characterization of HES-2, a mammalian helix-loop-helix factor structurally related to Drosophila hairy and enhancer of split.
RT Bur. J. Biochem. 215:645-652(1993).
RL -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH PROTEIN FOR THEIR TRANSCRIPTION.
CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX WITH A CO-REPRESSOR PROTEIN (GROUCHO).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG).
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY HAIRY-RELATED PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: D14029; BAA03118.1; -
PIR: S35037; S35037.
TRANSFAC: T01650; -
InterPro: IPR003015; HLH_MYC.
InterPro: IPR001092; HLH_dlm.
InterPro: IPR003650; Orange.
Pfam: PF00010; HLH; 1.
SMART: SM00353; HLH; 1.
SMART: SM00511; ORANGE; 1.
PROSITE: PS00038; HELIX_LOOP_HELIX; 1.
Transcription regulation; DNA-binding; Nuclear protein; Repressor.
DNA_BIND 13 26
FT DOMAIN 139 148
FT DOMAIN 154 157
WRPW MOTIF (REQUIRED FOR ACTIVITY)
(BY SIMILARITY).
SEQUENCE 157 AA: 17028 MW: B5D621E814AE0369 CRC64:

Query Match 6.9%; Score 6; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 KRRRAR 68
|||||
DB 23 KRRRAR 28

RESULT 12

HES2_HUMAN STANDARD; PRT: 173 AA.
AC Q9Y543; Q9Y542;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor HES-2 (Hairy and enhancer of split 2).
HES2.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH PROTEIN FOR THEIR TRANSCRIPTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG).
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY HAIRY-RELATED PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: AL031848: CAB46198.1: -
 DR EMBL: AL031848: CAB46199.1: -
 DR InterPro: IPR003015: HLM_Myc.
 DR InterPro: IPR003015: HLM_dim.
 DR InterPro: IPR003650: Orange.
 DR Pfam: PF00010: HLM: 1.
 DR SMART: SM00353: HLM: 1.
 DR SMART: SM00511: ORANGE: 1.
 DR PROSITE: PS00038: HELIX_LOOP_HELIX: 1.
 DR Transcription regulation: DNA-binding; Nuclear protein; Repressor:
 KW Alternative splicing: 26
 FT DNA_BIND 13 BASIC DOMAIN
 FT DOMAIN 27 71 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY)
 FT DOMAIN 139 167 PRO-RICH
 FT DOMAIN 170 173 WRPW MOTIF (REQUIRED FOR ACTIVITY)
 FT VARSPLIC 48 173 (BY SIMILARITY)
 FT NSMSKLEKADVLEMTVREFQELPASPMPAPLPDSTYRE
 FT GYSACVRLARVLPACRVLEPAPVAPRLLEHMRRASTGLD
 FT GGRAGDSSGSPAPAPAPAPAPAPAPAPAPAPSPSPCGGLW
 FT RFW -> DASGWHLPPLHAGNCFLYIQAPQPPA (IN
 ISOFORM 2).
 FT SO SEQUENCE 173 AA: 18470 MW: B3751F4576E84F3D CRC64:
 Query Match 6.9%; Score 6; DB 1; Length 173;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 63 KRRRAR 68
 DB 23 KRRRAR 28
 RESULT 13
 YE89_HAEIN STANDARD: PRT: 184 AA.
 AC P44215: 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein H11489.
 GN H11489.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OC NCBI_TaxID=727:
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=RD / KW20 / ATCC 51907;
 RC MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 Weidman J.F., Phillips C.A., Spillig T., Hedblom E., Cotton M.D.,
 Uterback T.R., Hanna M.C., Nguyen D.T., Sandek D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,
 Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus
 RT Influenzae Rd.";
 RT Science 269:496-512(1995).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: U32826: AAC23143.1: -
 DR TIGR: H11489: -

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 184 AA: 20782 MW: C298357DBED80211 CRC64:
 Query Match 6.9%; Score 6; DB 1; Length 184;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 79 PCSLL 84
 DB 73 PCSLL 78
 RESULT 14
 TM22_HUMAN STANDARD: PRT: 194 AA.
 ID TM22_HUMAN
 AC Q9Y584: 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitochondrial import inner membrane translocase subunit TIM22.
 GN TIM22 OR TIM22.
 GN Homo sapiens (human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606:
 OX [1]
 RN SEQUENCE FROM N.A.
 RP Bauer M.F., Brunner M., Neupert W., Gerdtitz K.-D., Hofmann S.;
 RA "Cloning and mapping of human Tim22, homologous to S.cerevisiae Tim22,
 RT a membrane-integrated component of the Tim22-54 machinery for the
 RT import of mitochondrial carrier proteins.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF MITOCHONDRIAL INNER MEMBRANE
 CC PROTEIN IMPORT PATHWAY. INVOLVED IN THE IMPORT OF PROTEINS
 CC TARGETED TO THE INNER MEMBRANE OF THE MITOCHONDRION SUCH AS
 CC MITOCHONDRIAL CARRIER FAMILY MEMBERS (BY SIMILARITY)
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TIM17/TIM22/TIM23 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF155330: AAD40106.1: -
 DR InterPro: IPR003397: TIM17.
 DR Pfam: PF02466: Tim17: 1
 KW Protein transport; Transmembrane; Mitochondrion; Inner membrane.
 FT TRANSMEM 74 94 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 170 190 POTENTIAL.
 FT SEQUENCE 194 AA: 20270 MW: DC986E23DF18377 CRC64:
 Query Match 6.9%; Score 6; DB 1; Length 194;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 81 SLLGY 86
 DB 26 SLLGY 31
 RESULT 15
 NRTN_MOUSE STANDARD: PRT: 195 AA.
 ID NRTN_MOUSE
 AC P97463: 01-NOV-1997 (Rel. 35, Created)
 DT

```

DT 01-NOV-1997 (Rel. 35, last sequence update)
DE 16-OCT-2001 (Rel. 40, last annotation update)
DE Neurturin precursor.
GN NRTN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 AND
RP 181-190.
RX MEDLINE=971100947; PubMed=8945474;
RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
RA Crendon D.J., Johnson E.M. Jr., Milbrandt J.;
RT "Neurturin, a relative of glial-cell-line-derived neurotrophic
RT factor ";
RL Nature 384:467-470(1996)
CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
CC HAEMOPOIETIC CELLS.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U78109; AAC52954.1; -.
DR HSP: 007731; IAGQ.
DR MGD: MGI:108417; Nrtu.
DR InterPro: IPR002400; GF_cysknol.
DR InterPro: IPR001839; TGF-beta.
DR Pfam: PF00019; TGF-beta; 1.
DR PRINTS: PR00438; GFCSKNOT.
DR SMART: SM00204; TCFB; 1.
DR PROSITE: PS00250; TGF-BETA_1; FALSE_NEG.
KW Growth factor; Signal.
FT SIGNAL 1 19
FT PROPEP 20 95
FT CHAIN 96 195
FT DISULFID 101 163
FT DISULFID 128 192
FT DISULFID 132 194
FT DISULFID 162 162
SQ SEQUENCE 195 AA; 22219 MW; ABE21BB35D417448 CRC64;

```

```

Query Match 6.9%; Score 6; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 64 RRRAP 69
DB 91 RRRAP 96

```

Search completed: June 21, 2002, 08:25:25
Job time: 310 sec